详细解答

STEP I 1993，第 14 题

概率与统计 · 指数分布的等候时间

STEP I 1993，第 14 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/93-s1-q14

关键思路

重要的不是蟾蜍先生的速率，而是他走完最后一公里所需的时间 T，而 $T = 1/V$。V 的密度函数看来不好处理，但它的累积分布函数（cumulative distribution function）却非常简单；把它转化为关于 T 的叙述，便可看出 T 服从参数为 α 的指数分布（exponential distribution）。此后，每个绿灯时段的概率都是两个指数函数之差，而把所有绿灯时段加起来，就得到一个等比级数（geometric series）。

解答

到达时间的分布

设 V 为速率，单位为公里每分钟，并设 $T = 1/V$ 为他到达交通灯的时间，以分钟为单位，由交通灯转绿的一刻起计。留意到

$$\frac{\mathrm{d}}{\mathrm{d}v}\,\mathrm{e}^{-\alpha/v} = \alpha v^{-2}\mathrm{e}^{-\alpha/v},$$

以及当 $v \to 0^+$ 时 $\mathrm{e}^{-\alpha/v} \to 0$，可知 V 的分布函数为

$$\mathrm{P}(V \le v) = \int_0^v \alpha u^{-2}\mathrm{e}^{-\alpha/u}\,\mathrm{d}u = \mathrm{e}^{-\alpha/v} \qquad (v > 0).$$

（当 $v \to \infty$ 时它趋向 1，所以这确实是一个密度函数。）由于 $V > 0$，对 $t > 0$ 我们有 $T > t \iff 1/V > t \iff V < 1/t$，所以

$$\mathrm{P}(T > t) = \mathrm{P}\left(V < \tfrac{1}{t}\right) = \mathrm{e}^{-\alpha t}.$$

因此 T 服从参数为 α 的指数分布。

想一想 为甚么要用 T 而不用 V 来计算?

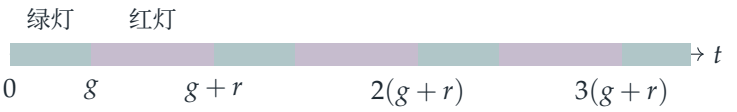
交通灯是以时间来描述的，所以我们关心的每个事件，都是 T 的取值所在的一个区间。若要直接把 V 的密度函数换算成 T 的密度函数，便须作变量代换（change of variable）；改为经由 $\mathrm{P}(T > t)$ 入手，则可避开这一步，一行便得出答案。

一个绿灯时段

在时间 0，交通灯转为绿灯。绿灯时段为 $[0, g)$，红灯时段为 $[g, g+r)$，然后绿灯时段为 $[g+r, (g+r)+g)$，如此类推。第 $(n+1)$ 个绿灯时段是 $n(g+r) \le t < n(g+r)+g$。由于 T 是连续的，是否包括端点并无分别，而

$$\begin{aligned}\mathrm{P}(n(g+r) < T < n(g+r)+g) &= \mathrm{P}(T > n(g+r)) - \mathrm{P}(T > n(g+r)+g)\\ &= \mathrm{e}^{-\alpha n(g+r)} - \mathrm{e}^{-\alpha(n(g+r)+g)},\end{aligned}$$

正如题目所求。（原试卷把第二个指数中的右括号印错了位置；该指数应为 $-\alpha(n(g+r)+g)$，如上所示。）以上推算对 $n = 0$ 同样成立，那就是第一个绿灯时段，此时交通灯还未曾转过红灯。



遇上绿灯的概率

他在绿灯时通过，当且仅当 T 落在其中一个绿灯时段之内。这些时段互不相交（disjoint），所以我们把它们的概率相加，包括 $n = 0$:

$$\begin{aligned}\mathrm{P}(\alpha) &= \sum_{n=0}^{\infty}\left(\mathrm{e}^{-\alpha n(g+r)} - \mathrm{e}^{-\alpha n(g+r)}\mathrm{e}^{-\alpha g}\right) = (1-\mathrm{e}^{-\alpha g})\sum_{n=0}^{\infty}(\mathrm{e}^{-\alpha(g+r)})^n\\ &= \frac{1-\mathrm{e}^{-\alpha g}}{1-\mathrm{e}^{-\alpha(g+r)}},\end{aligned}$$

这里求了等比级数之和；该级数收敛（converges），因为 $0 < \mathrm{e}^{-\alpha(g+r)} < 1$（由于 $\alpha > 0$ 及 $g+r > 0$）。

极限

当 $\alpha \to \infty$ 时，$\mathrm{e}^{-\alpha g} \to 0$ 和 $\mathrm{e}^{-\alpha(g+r)} \to 0$ 同时成立（因为 $g > 0, g+r > 0$），所以

$$\mathrm{P}(\alpha) \to \frac{1-0}{1-0} = 1.$$

这是合理的：α 很大表示蟾蜍先生开得非常快，几乎每次都能在第一个绿灯时段结束前通过。

当 $\alpha \to 0$ 时，分子和分母都趋向 0，所以要更小心处理。在分子和分母中分别写成 $1-\mathrm{e}^{-x} = \mathrm{e}^{-x}(\mathrm{e}^{x}-1)$：

$$\mathrm{P}(\alpha) = \frac{\mathrm{e}^{-\alpha g}\left(\mathrm{e}^{\alpha g}-1\right)}{\mathrm{e}^{-\alpha(g+r)}\left(\mathrm{e}^{\alpha(g+r)}-1\right)} = \mathrm{e}^{\alpha r}\cdot \frac{\dfrac{\mathrm{e}^{\alpha g}-1}{\alpha g}}{\dfrac{\mathrm{e}^{\alpha(g+r)}-1}{\alpha(g+r)}}\cdot \frac{g}{g+r},$$

其中我们把分子和分母同除以 α，并补上因子 g 和 $g+r$。当 $\alpha \to 0$ 时，$\mathrm{e}^{\alpha r} \to 1$，而且 $\alpha g \to 0$ 和 $\alpha(g+r) \to 0$ 同时成立，所以由 $(\mathrm{e}^{x}-1)/x \to 1$ 可知两个分式各自趋向 1。因此

$$\mathrm{P}(\alpha) \to \frac{g}{r+g} \qquad \text{当} \alpha \to 0.$$

这时蟾蜍先生开得非常慢，他的到达时间分散在很多个周期之中，所以遇上绿灯的机会就只是每个周期中绿灯所占的比例。

验算

- 取 $\alpha = 0.7, g = 1, r = 2$，公式给出 $\mathrm{P}(\alpha) \approx 0.5737$。模拟 4×10^{5} 次车程，以求 $\mathrm{e}^{-\alpha/v}$ 的反函数的方法对 V 抽样，结果为 0.5744，而 $n = 0, 1, 2$ 各个绿灯时段的概率亦一一吻合。
- 取 $g = 1, r = 2$：$\mathrm{P}(10^{-4}) \approx 0.33337$，接近 $\frac{1}{3}$，而 $\mathrm{P}(10) \approx 0.99995$。
- 若 $r = 0$，交通灯永远是绿灯，而公式对每个 α 都给出 $\mathrm{P}(\alpha) = 1$，正该如此。

常见错误

- 漏掉 $n = 0$。第一个绿灯时段在任何红灯之前，所占的概率最大。
- 照抄试卷上印的指数 $-\alpha(n(g+r))+g$，那是 $-\alpha\big(n(g+r)+g\big)$ 的误印。
- 把 V 本身当作到达时间，或者尝试求 $\mathrm{E}(T)$。事实上这里的 $\mathrm{E}(V)$ 是无限大，但题目完全用不着它。

STEP 之外

通过分布函数把 V 的分布转化为 $T = g(V)$ 的分布，是变换随机变量（random variable）的标准方法。这里 V 服从逆指数分布（inverse-exponential distribution）。极限 $\alpha \to 0$ 是

一个普遍想法的小例子：一个分散在许多个周期之中的随机时间，落在各个阶段的机会与该阶段的长度成正比，这正是更新理论（renewal theory）的起点。Grimmett 与 Welsh 的 *Probability: An Introduction*（概率论入门）对随机变量的变换和指数分布都有论述。

Beyond STEP 详细解答：https://beyondstep.keithmcnulty.org/zh-hans/。解答文字采用 CC0 1.0；题目本身版权属 Cambridge University Press & Assessment 所有（©），本文件不予转载。由机器从英文翻译而成；尚未经审阅。